Stream Restoration in the Pacific Northwest: Analysis of Interviews with Project Managers

Jeanne M. Rumps,[1,2] Stephen L. Katz,[3] Katie Barnas,[3] Mark D. Morehead,[1] Robin Jenkinson,[1] Stephen R. Clayton,[1] and Peter Goodwin[1]

Abstract

Hundreds of millions of dollars per year are spent on river restoration in the Pacific Northwest (PNW), but little is known about the effectiveness of this effort. To help address this gap, we analyzed a database containing 23,000 projects at 35,000 locations in the region. We selected a subset of these projects for interviews using a survey instrument developed by a national team of scientists. In total, 47 project contacts in the PNW were interviewed to learn from the individuals directly involved in restoration. At least one-third of the projects surveyed (34%) did not conduct sufficient monitoring to evaluate effectiveness. More than two-thirds (70%) of all respondents reported their projects were successful, but 43% either have no success criteria or are unaware of any criteria for their project. Although almost two-thirds (66%) of respondents anticipate a need for ongoing project maintenance, less than half (43%) have maintenance funds available. These findings suggest that establishing a connection between effectiveness monitoring and project implementation is not a usual component of project design. Consequently, we can only assess the benefits in a few isolated projects and cannot quantify the cumulative benefits of restoration on a larger scale. These findings highlight the need for (1) planning prior to implementation of restoration projects that accounts for monitoring design; (2) coordinated effectiveness monitoring to assess cumulative effects of restoration; and (3) management and maintenance of projects based on real measures of project performance.

Key words: effectiveness monitoring, implementation monitoring, Pacific Northwest, restoration.

Introduction

In the Pacific Northwest (PNW), restoration of freshwater habitat forms a cornerstone of both conservation and management strategies. Between 2001 and 2003, nearly $400 million federal dollars per year were spent on management activities aimed at rehabilitating the Columbia River Basin alone (GAO 2002; Roni et al. 2002; Katz et al. 2007). Nationally, freshwater restoration expenses extend to the billions of dollars (Lavendel 2002; Malakoff 2004; Bernhardt et al. 2005). Given this large investment in freshwater habitat improvement, how successful have the restoration actions been? Unfortunately, there is little pre- and postproject information explicitly linking this restoration with habitat responses (NRC 1992; Bash & Ryan 2002; GAO 2002). Indeed, although the cumulative effect of all habitat projects has a detectable impact, there is insufficient information to determine what kind of projects are effective and under what circumstances (Paulsen & Fisher 2005). Important steps in defining restoration success include both complete documentation of restoration actions via implementation monitoring and establishing a functional connection between actions and ecological responses via effectiveness monitoring (NRC 1992; Kondolf & Micheli 1995; Kondolf 1998; Palmer 2005; Katz et al. 2007).

Implementation monitoring assesses whether planned activities, documented in project-tracking data systems, were executed as designed (MacDonald et al. 1991). Over the past 3 years, the National Marine Fisheries Service has compiled habitat restoration information within a project-tracking database (PNW Salmon Habitat Restoration Project, "PNW Database"). The PNW Database currently contains implementation information on over 23,000 restoration projects in Oregon, Washington, Montana, and Idaho. In a related database compilation effort, the National River Restoration Science Synthesis (NRRSS) documented river and stream restoration activities on a national scale (Bernhardt et al. 2005, 2007). When combined, a national database ("NRRSS Summary Database") was created that contains information on approximately 40,000 river restoration projects.

Although the PNW Database provides abundant information on planned projects (who, what, where, and when), the ecological and biological effects of individual projects on target species remain largely unexplored with adequate effectiveness monitoring. There are diverse definitions for effectiveness monitoring, both at the project scale and at the watershed scale. The components of adequate effectiveness monitoring include clearly articulated

[1] Center for Ecohydraulics Research, University of Idaho, 322 East Front Street, Suite 340, Boise, ID 83702, U.S.A.
[2] Address correspondence to J. M. Rumps, email jrumps@idfg.idaho.gov
[3] Northwest Fisheries Science Center, NOAA Fisheries, 2725 Montlake Boulevard East, Seattle, WA 98112, U.S.A.

questions to define the scale of inference, pre- and post-implementation data collection with replication and reference sites, and characterization of data quality and performance (e.g., Green 1979; MacDonald et al. 1991; Thompson et al. 1998). Because of the lack of funding and institutional mechanisms, restoration project monitoring is often performed opportunistically or based on single, isolated efforts. Monitoring rarely includes long-term commitments with sufficient spatial and temporal contrasts to determine the impact of restoration actions on watershed processes (Reeves 1991*a*, 1991*b*; NRC 1992; Downes et al. 2002; Sutherland et al. 2004). Documenting responses is challenging as physical and biological responses to restoration can be highly variable over a range of scales (Kondolf & Micheli 1995; Frissell & Ralph 1998; Kondolf et al. 2001; Roni et al. 2002).

Defining what constitutes successful restoration is contextual; project sponsors have diverse intents for the restoration actions they undertake (Palmer et al. 2005). In the PNW, there are several large-scale initiatives that motivate restoration including the Endangered Species Act (ESA), the Clean Water Act, the Northwest Power Planning Act, and numerous nonregulatory programs. Each of these initiatives provides its own specific context for defining what makes restoration successful.

Given the lack of existing monitoring design and the paucity of pre-treatment monitoring of restoration projects, alternative methods to assess project success must be developed. In the absence of standardized implementation or effectiveness monitoring protocols, restoration project managers may offer the next best available knowledge regarding project implementation and success. To assess the success of restoration PNW projects, we surveyed 47 project sponsors and contacts as a subsample of projects in the PNW Database. This survey was one component of a national survey of 317 stream restoration project managers from seven regions throughout the United States (Bernhardt et al. 2007). In this article, we present phone survey results for the PNW and applicable lessons learned from practitioners. Survey results were compared with a random set of responses to evaluate survey reliability and aid in interpretation. We also describe how managers, engineers, scientists, and ecologists can use this information to improve the science of restoration activities in achieving cumulative watershed benefits.

Methods

Phone Survey Design

The PNW Database, with restoration projects located in Oregon (~50%), Washington (~30%), Idaho, and Montana served as the sampling universe (Katz et al.). Projects were eligible candidates for an interview if they met three criteria: (1) were implemented or completed between 1996 and 2002; (2) had information about a project contact (e.g., an individual or agency name); and (3) listed at least one of four selected project goals (riparian management, water quality management, in-stream habitat improvement, or channel reconfiguration). We sampled 12 projects per intent to maintain consistency with the coordinated national effort (Bernhardt et al. 2007). Noninclusion in the survey (Table 1) occurred when multiple attempts to contact the project manager were unsuccessful or if the contact had previously been interviewed regarding another project. If an interview request was declined or the project abandoned, we sampled the next project on the randomly ordered project list. In the PNW, only 11 water quality improvement projects were surveyed because of sample nonresponses, making the total sample population 47.

The Survey Process

The phone survey was developed as part of a larger national effort to assess restoration project success in the absence of large-scale effectiveness monitoring protocols (Bash & Ryan 2002; Bernhardt et al. 2005). The questionnaire is described in detail in Bernhardt et al. (2007) and therefore only briefly described here.

Willing interviewees were provided contextual material, including a list of survey topics. Each interviewee was asked the standard set of 47 questions as well as several open-ended questions to elicit narrative on additional lessons learned in performing restoration actions. The duration of the recorded phone interviews varied between 30 and 90 minutes.

After the interviews were complete, we developed a semiquantitative assessment of relative project success by attaching numeric values to the answers to 17 of 47 total questions in the survey. The basis of this was a scheme for evaluating project success that included five elements (Palmer et al. 2005): (1) articulation of the desired state (=*guiding image*); (2) measurable improvement in ecological conditions (=*ecosystem improved*); (3) self-sustenance after the restoration (=*resilience*); (4) implementation not inflicting irreparable harm (=*no lasting harm*); and (5) rigorous effectiveness monitoring to allow performance assessments (=*ecological assessment*).

Table 1. Primary reasons for noninclusion of interviews in sampling.

Reasons for declining interview	*# of answers*
Not stream restoration	11
Interviewee declined	4
Interviewee already conducting interview on separate project	2
Project does not exist	1
In litigation and cannot discuss	1
No contact person or project information found	1
Unknown reason	1
Total count	21

We assessed three of the five elements (*guiding image*, *ecosystem improvement*, *and ecological assessment*) by summing the scores to the subsets of the 17 questions that addressed those measures of project success. The scores in each of these three categories were then summed to provide a *total score* to allow a comparison of overall performance (see online accessory material).

Characterization of the Survey Performance

Selecting a set of questions and assigning a scoring system based on potential answers a priori is not a statistical test. Rather it is an expectation, and it is difficult to attach importance to the answers interviewees might give. There are no preexisting, objective, and independent criteria that define a "good" or "bad" score, nor is there a preexisting expectation for the distribution of scores in the survey population. One could set subjective thresholds that define "good," "intermediate," and "bad," scores, but proving these are not arbitrary is challenging. To provide objective basis for discriminating high- from low-performance projects, we compared survey results to an independent, randomly generated population of results. This provides a null hypothesis that survey answers are no different from randomly generated answers. This null hypothesis tests the survey itself; the hypothesis will not be rejected unless the projects are perceived as high performance and the survey detects that perception.

To generate the independent test population, randomly generated integer scores were simulated for the 17 questions in the survey characterization. Multiple-choice answers were given an equal probability of being selected. In two cases (questions 28 and 29), answers were a list of items with no upper limit. For these questions, the random answers were drawn from a distribution approximating an exponential decay function (i.e., many low scores, few high scores).

We generated 7,000 simulated survey responses to the relevant questions and assembled the four category distributions (*guiding image, ecosystem assessment, ecosystem improvement*, and *total score*). To identify restoration projects perceived as performing "poorly" or "well" relative to the random distribution, we generated 95% confidence intervals for observed distributions with a bootstrap (Efron & LaPage 1992). Five thousand responses from four category distributions were randomly subsampled with replacement from the 7,000 simulations, and the 0.025 and 0.975 percentiles were identified within the subsample to form a single bootstrap draw. Bootstrap sample size was a compromise between analytical expectation of an infinite sample size and computational efficiency. With the available range of discrete outputs, mean square errors asymptotically converged well below 5,000 draws (Hall 1992). The mean values for the 0.025 and 0.975 percentiles were estimated and resolved to the next lower or higher integer, respectively. Projects that scored below or above the 95% confidence interval were viewed (according to respondents' answers) as performing "poorly" or "well," respectively. Project scores that fell within the confidence intervals would have been observed frequently by chance and were not deemed either high- or low-performance projects.

Differences between the survey results and the randomly generated distributions were tested with a two-sample Kolmogorov–Smirnov test (Zar 1984). Correlations of scores between categories were assessed with a Kendall's τ rank-order correlation coefficient (Zar 1984). Estimates of τ were made with SYSTAT 11 statistical software (SYSTAT Software, Inc., Richmond, CA, U.S.A.). Critical significance values for τ were two tailed given we had no expectation that project assessments would be positively or negatively correlated.

Results

The analysis of our phone interviews attempts to indicate why projects occurred, whether the results were perceived as successful by those involved with the project, and how success was determined. Absent targeted monitoring, this approach offers a subjective means for determining the ecological benefits or effects of restoration actions.

Respondent Characteristics

Half of all individuals interviewed represented federal agencies (Fig. 1); the other half represented state and local government agencies, tribes, nongovernmental organizations, or the timber industry. The interviewees showed a diverse connection to projects with approximately one-third identifying themselves as a manager or coordinator and, thus, the person most knowledgeable of the project (Fig. 2). The remaining proportion of respondent roles included designer, implementer, evaluator, consultant, or funder.

Project Types and Motivation

As noted previously, only projects with primary intents of water quality improvement, in-stream habitat improvement, channel reconfiguration, and riparian management were eligible for interview selection. In most cases, the original classification of primary intent was based on the activity undertaken as reported in the PNW Database (Katz et al. 2007). However, interviewees often indicated that the actual primary intent differed from the four categories (Fig. 3), revealing a discrepancy between project reporting and project sponsor perception. This discrepancy exists in 16 of 47 projects interviewed, largely because of the multiple (and sometime unintended) effects of a single activity. Thus, a project may have been selected for an interview based on the four strata and did indeed implement the described activity; however, the project contact may have viewed the project as principally addressing an alternative intent.

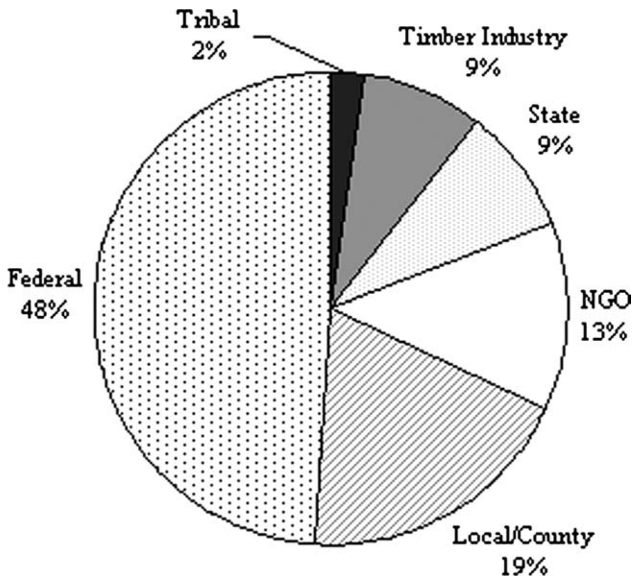

Figure 1. Distribution of interviewee affiliation ($n = 47$).

When asked why the restoration action was undertaken at that time and presented with a multiple-choice answer, the most frequent response was to address the greatest habitat degradation factor, followed by the presence of protected fish (Fig. 4). When asked why the restoration was performed at this particular location, the most frequent response cited reasons of ecological concern (Fig. 5). "Ecological concern" was substantially more common than the alternatives, seven of which were closely grouped between "protected fish presence" and "infrastructure concerns."

Project Design and Maintenance

Several questions were designed to elicit information on the scale of decision-making and degree of coordination with regional planning efforts during restoration. Over half (53%) of respondents reported their project was an element of a larger watershed management plan indicative of a guiding image for the basin. In a related question, 77% of interviewees responded that their project activities were specifically linked to other implemented or planned restoration in the river segment. Replies to both questions indicate that the majority are part of a larger, coordinated effort and suggest the presence of a guiding image. However, the survey did not provide a mechanism to assess the temporal schedules of projects; projects implemented 10 years apart with parallel intents may, in fact, be just coincidental.

When questioned about the sources of knowledge used to create, implement, and/or evaluate design plans, almost half of respondents (49%) indicated using professional expertise or an "expert" opinion (e.g., hydrologist, biologist, engineer, geomorphologist, or ecologist). The second most used resource was government guidelines (manual, book, or agency report) and prior experience (both at 22%, Table 2).

To understand whether or not adaptive management was being incorporated, we asked if the interviewee anticipated a need for project maintenance. The majority (66%) of respondents indicated that ongoing maintenance would be needed to ensure long-term benefits. However, only 43% of subjects had funding allocated for maintenance, which indicates a de facto short-term project lifetime in many project design schemes.

Project Monitoring

The initial question regarding project monitoring was simply: "Did you collect monitoring data specific to this project?" Absent a regional standard for what constitutes "implementation monitoring" or "effectiveness monitoring," answers to this question reflect a diversity of definitions in use. Eighty-one percent of respondents reported collection of monitoring data. Among those, 70% monitored physical, biological, or chemical (water quality) data (Fig. 6). When asked about the duration of monitoring, 18% of respondents reported a single observation (i.e., monitoring for implementation only), suggesting almost one in five did not discriminate between effectiveness and implementation monitoring. Eight percent of those reporting monitoring characterized it as "physical monitoring" with a "one-shot deal" duration (Fig. 7, 43% of 18%). Further questioning revealed that this commonly amounted to referencing preimplementation topographic survey data in the design phases. An additional 2% of respondents completed just one visual assessment to ensure implementation of the project and believed this to be monitoring (Fig. 7). A striking feature is that 23% plan

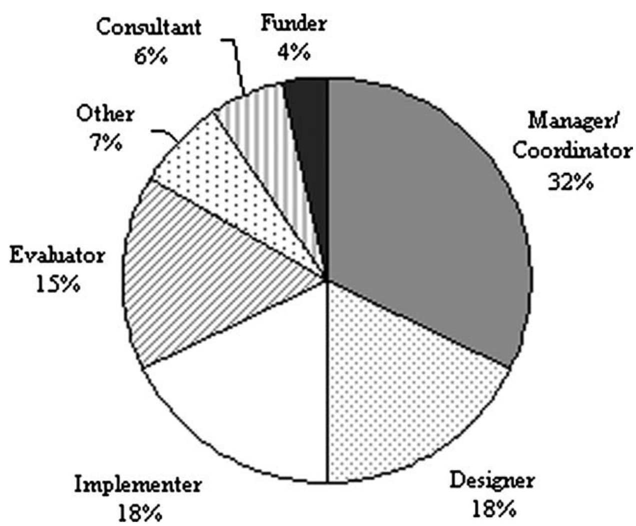

Figure 2. Role of interviewees in regard to their particular restoration project.

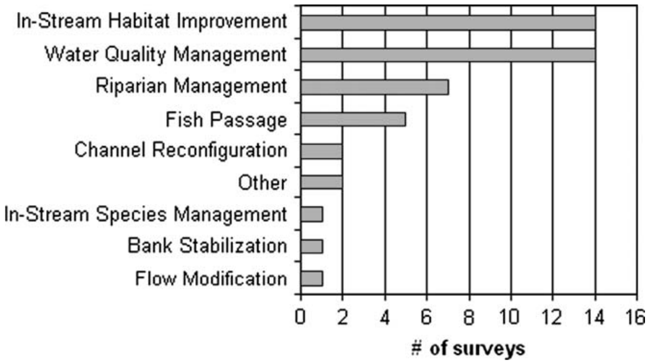

Figure 3. Distribution of restoration project intentions based on phone interviewee responses ($n = 47$).

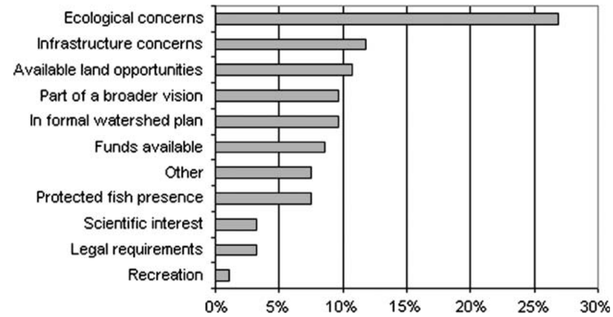

Figure 5. Distribution of why the particular location was chosen for the restoration work over other possible locations.

to monitor for more than 10 years. Of the multiple monitoring efforts, the majority reported an annual frequency.

If we assume the one-time monitoring respondents ($n = 7$) were solely to ensure project implementation and assume those who monitored 2 years or more ($n = 31$) were monitoring for some measure of effectiveness, we can estimate that 66% of all subjects performed some version of effectiveness monitoring.

Project Evaluation

When asked if specific success criteria were stated in the design plan, the majority responded yes (57%), 32% replied no, and 11% did not know. Of the 57% that answered yes, 38% stated the project has *completely* met the criteria, 38% stated the project has *partially* met the criteria, and 23% indicated that it is too soon to tell. Thus, projects that met preimplementation performance criteria either partially or completely amounted to 43% of the survey population.

In a separate question aimed toward qualitative evaluation, we asked whether the subject personally considered the project a success (either socially or environmentally). When presented with multiple-choice answers, 71% indicated the project was a *complete* success, 23% said a *partial* success, 4% said it is too soon to tell, and 2% stated the project was not at all successful (Fig. 8). However, only 12% of respondents based their answers on actual measurements (Fig. 9). For those that indicated less than a complete success, the most often stated reasons included inadequate design or the low survival rate of riparian plantings.

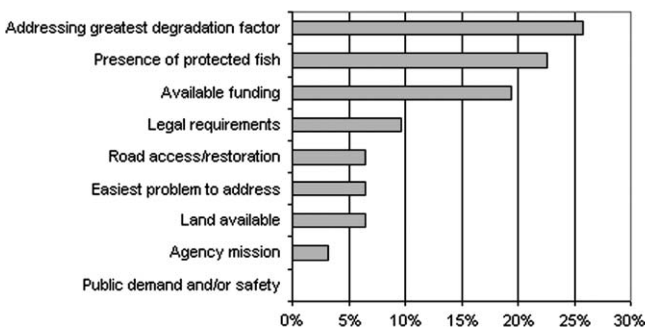

Figure 4. Distribution of factors that motivated the restoration project based on phone interviewee responses.

Lessons Learned

When asked if they would make changes in retrospect, 64% of interviewees indicated that they would change how the project was monitored and evaluated. The most common reason for wanting more and better monitoring data was to scientifically evaluate project effectiveness, a common avenue to leverage future funding sources. Other desired changes included having more funds allocated toward methodical monitoring to enable documentation of project effectiveness.

Scoring for Ecological Success

The distribution of interview scores in the *guiding image* category differed significantly from the randomly generated distribution (K–S $D = 0.35$, $p < 0.02$; Fig. 10a). The interview scores showed a relatively uniform distribution over the range of observed values (0–13, Fig. 10a), whereas the simulated distribution had a central tendency with a large number of high scores. For *guiding image*, the frequency of high values results from the incorporation of question 37, which has three possible answers, assigned

Table 2. Source of knowledge used for planning and design of restoration project.*

	# of answers	%
Individual expertise	63	49
Agency guidelines	29	22
Past experience	28	22
Models or project site analysis	5	4
Peer-reviewed journal article	2	2
Workshop or short course	2	2
	129	100

* Respondents were not limited to one answer.

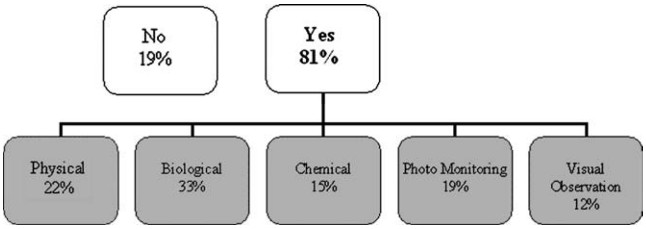

Figure 6. Distribution of answers to questions pertaining to monitoring; specifically, "Did you monitor," and, if yes, "What was monitored?"

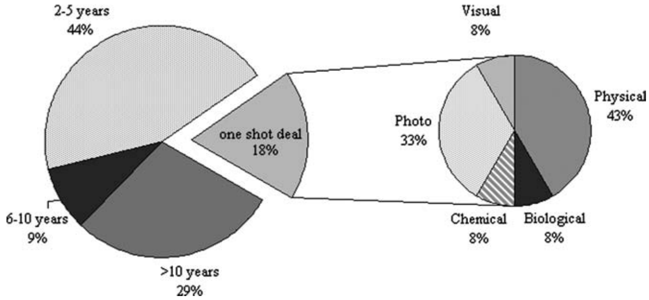

Figure 7. Distribution of interviewee answers regarding the duration of intended monitoring.

numerical scores of 1, 2, or 6. In spite of this positive bias in randomly generated answers, the distribution of interview responses shows that none were viewed as performing well with respect to *guiding image* (score >14). In contrast, almost 20% of the interviews were viewed as performing poorly in this category (score <2).

The interview scores for *ecological assessment* have a relatively uniform distribution with no clear central tendency (Fig. 10b). The simulated distribution, in contrast, resembles a Poisson process with a clear peak and a long tail out to extreme high values, and it differed significantly from the survey results (K–S $D = 0.22$, $p < 0.02$). As discussed above, this is a result of unbounded questions 28 and 29. There were only six projects (13%) that scored high, but nine projects (18%) that scored low in this category. Readers are cautioned to remember that variation of project complexity is high, and all projects do not require equal and intensive assessment. The statistical weighting of this question does not allow the necessary case-by-case analysis with respect to project type (e.g., riparian plantings may be sufficiently assessed by periodic visual inspections; Palmer et al. 2005).

The distribution of *ecosystem improvement* interview scores was not different from the randomly generated distribution (K–S $D = 0.18$, $p > 0.20$; Fig. 10c). Both distributions display a central tendency, and the mean of the randomly generated scores was similar to the interview scores. In this category, there were no scores less than the lower confidence interval. One single project was viewed as performing well with respect to *ecosystem improvement* (score >12).

In the *total score* category, the interview scores showed no central tendency, whereas the simulated scores show a clear central tendency and an approximate symmetrical distribution. This difference was significant (K–S $D = 0.25$, $p < 0.001$) (Fig. 10d). Because the scores are cumulative, it is clear that the extremely high-performing projects with respect to *ecological assessment* will also perform well in *total score*. The range of total scores in the PNW was 0–56 compared with national values that went as high as 78 as a consequence of individual projects in other regions having elaborate monitoring designs (Bernhardt et al. 2007).

Similar to high-scoring projects, the 13 interviews (25%) that received low scores in one category are likely to perform poorly across other categories. Analysis of rank-order correlation of the scores in the four different categories reveals that performance across categories is indeed correlated. The largest rank-order correlation (Kendall's τ) was between total score and ecological assessment score at $\tau = 0.78$ ($p < 0.001$, two-tailed probability), and the lowest was between guiding image and ecological assessment at $\tau = 0.39$ ($p < 0.01$). This indicates that projects with a well-established guiding image are also well monitored and potentially improving the ecosystem. Likewise, projects that lack a clear guiding image are likely to be unsuccessful in other ways.

Discussion

The major objective of this study was to assess habitat restoration project performance in the PNW. The interviews allow us to elicit the opinions of those who worked closely on such projects, but their subjective expressions of project success do not replace rigorous effectiveness monitoring. Overall, interviewees were very positive about the outcomes of their activities even though less than half of those interviewed had used any preimplementation performance criteria.

Without adequate effectiveness monitoring, we cannot learn from our successes and failures, and the science of river restoration will not advance (Kondolf & Micheli 1995). Project evaluation as described in Palmer et al. (2005) defined the success of restoration projects by performance in several large-scale elements including guiding image, ecological assessment, and ecosystem improvement. By assigning a scoring component, we have "classified" individual projects based on interviewee response as it pertains to these three facets of performance. It is, however, more difficult to generalize to the populations of projects.

The observed distributions fell into two classes: *ecosystem improvement* (which resembled the simulated distributions) and those remaining. The survey results and the scoring system suggest that while some monitoring of restoration projects is common, it typically consists of implementation rather than effectiveness monitoring. By definition, assessments of ecosystem response to restoration are limited by the information quality obtained in

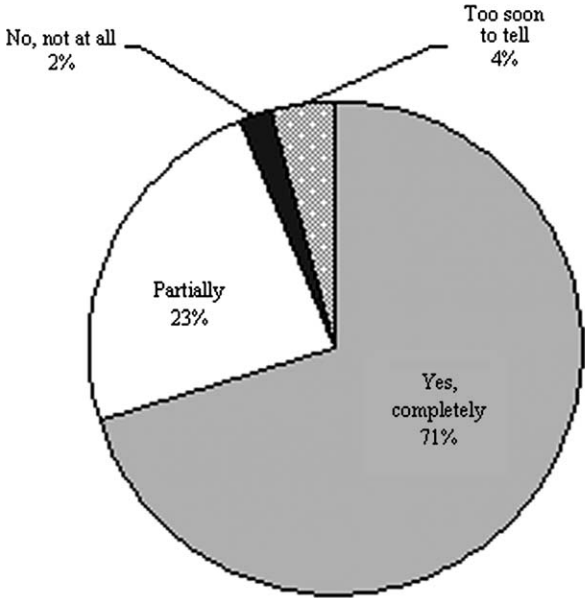

Figure 8. Personal evaluation of project success ("Do you consider this project successful?").

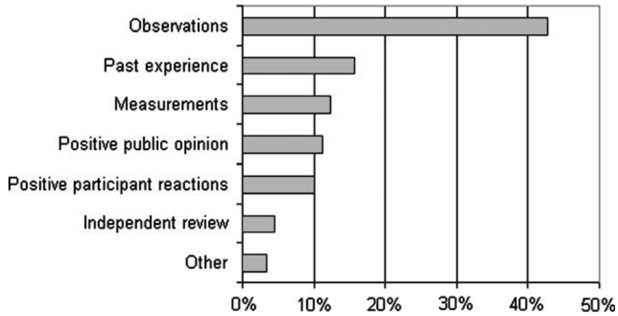

Figure 9. Basis of success, as provided by telephone interviewees.

monitoring (Sutherland et al. 2004); if the monitoring program is superficial, ecosystem assessments are limited to assumptions. With a paucity of adequate effectiveness monitoring in the survey population, we observed subjective answers indistinguishable from a random distribution.

Our comparison of distribution shapes also points to a limit in our simulation approach. By defining the simulated test distributions this way, we have set up the null hypothesis that survey answers are no different from those generated with a random process—effectively comparing our survey answers to noise. This design is robust for accurately detecting a good scoring project (next section). However, it may be more desirable to compare survey answers to an independent reference condition for "good" or "bad" projects. If one were to independently identify high-performance restoration actions whose scores could be characterized as "good," one could then determine how frequently a good score is observed within the simulated population. In this way, we could estimate the likelihood of failing to detect an existing good project (our type II error rate). However, setting these independent standards for good and bad projects objectively would require additional information on project design, monitoring, and impact than we currently possess.

Phone Interview Bias

The results indicate that the generally positive impressions of project respondents are largely subjective and potentially subject to hidden bias (Rosenbaum 2002). Interviewees were likely invested in project implementation, either in effort expended or in financially. This investment creates an obvious potential source of bias when surveyed regarding project success. For example, in the case of a project failure, an interviewee may be disinclined to participate in the survey at all, or, if willing to participate, disinclined to discuss and elaborate on project performance. If true, then we may expect to arrive at an overly optimistic assessment of restoration success.

Bias may also exist in a narrower context through the questions themselves. Although attempts were made to eliminate bias in the wording of the questionnaire, several questions were specifically designed to assess the quality of monitoring. If the interviewees suspected that monitoring was a priority for the interviewer, they may have biased their answers to highlight this.

There are additional potential sources of hidden bias in the source data itself. For example, the PNW Database is a sample and not a census. Although it does include a large number of projects (>23,000), its accuracy is a product of the quality of regional project-tracking data systems and the associated detective work in acquiring the data. As discussed in Katz et al. (2007), the PNW Database is likely an underestimate of the total projects in the region. Although the large sample size provides a measure of confidence in statistical inferences, readers need exercise appropriate caution.

Additionally, in the PNW, the ESA listings of five salmon species and the spotted owl are a highly significant influence shaping freshwater restoration (NRC 1996; GAO 2002; Ruckelshaus et al. 2002; Katz et al. 2007). Our analysis further demonstrates that the presence of protected species motivates restoration activity, linking the projects specifically to local priorities. This reflects that local management priorities may influence or bias the reporting of project activities (the source of the PNW Database).

The possible presence of bias in the survey design makes the simulation of random scores an important exercise. The distribution of random scores allows us to critique the interview process in retrospect. The results from the simulations indicate two important features of the survey: (1) a very high score is needed to be deemed high performance with respect to a random process and (2) the high standard reduces the potential impact of hidden bias. Given the small number of discrete scores that are possible in each category, the area outside of the 95% confidence interval will only contain the most extreme values regardless of the observed distribution. As it turns out, we observed low scores more frequently than high scores in spite of our suspicion that project representatives would be inclined to highlight the positive aspects of their projects during interviews. This suggests that while bias may be present, the survey instrument is robust to its influence. Readers are reminded that the cost of this robustness is the possibility that we will fail to recognize projects with good performance (see above).

Evaluation of Project Performance

As previously noted, a large majority of those interviewed (70%) reported complete success, but 43% either have no success criteria or do not know if any exist. It is unfortunate that there is little connection between the evaluation of success and the actual use of success criteria. Even among those that collected measurements, there is little or no experimental design employed to investigate causal linkages between restoration and ecological impact at the project or regional scale (design of monitoring programs is discussed elsewhere: e.g., Green 1979; Carpenter et al. 1989; Underwood 1994; Downes et al. 2002). Although interviewees indicated that multiple projects were indeed coordinated, support was lacking in the form of regional data collection, sharing, or analysis to allow synthesis of restoration impacts on large spatial scales. Where impacts of habitat restoration are examined on the scale of large watersheds, there have been cases where no effect could be traced to categories or groups of projects (Paulsen & Fisher 2005).

Understanding the ecological response to restoration requires project monitoring and larger-scale coordinated efforts than what are currently present. Indeed, interviewees in this study reported 66% of projects incorporated some form of effectiveness monitoring, but only 7% of project records in the PNW Database (Katz et al. 2007) and 10% nationally (Bernhardt et al. 2005) have associated monitoring. The survey responses suggest that monitoring is increasingly recognized as a critical component in validating success of restoration, but monitoring is not being shared and informing regional activities sensu Palmer et al. (2005). Additionally, 66% of project managers surveyed believe their project will need ongoing maintenance to persist, but only half have the necessary funding to cover maintenance. Although interviewees are aware of the fragility of their projects, associated monitoring and adaptive management are not designed to capture the maturation, resilience, and progressive evolution.

Throughout the PNW, there are thousands of well-intended restoration projects but few monitoring-based inferences of effectiveness. Given the impracticality of

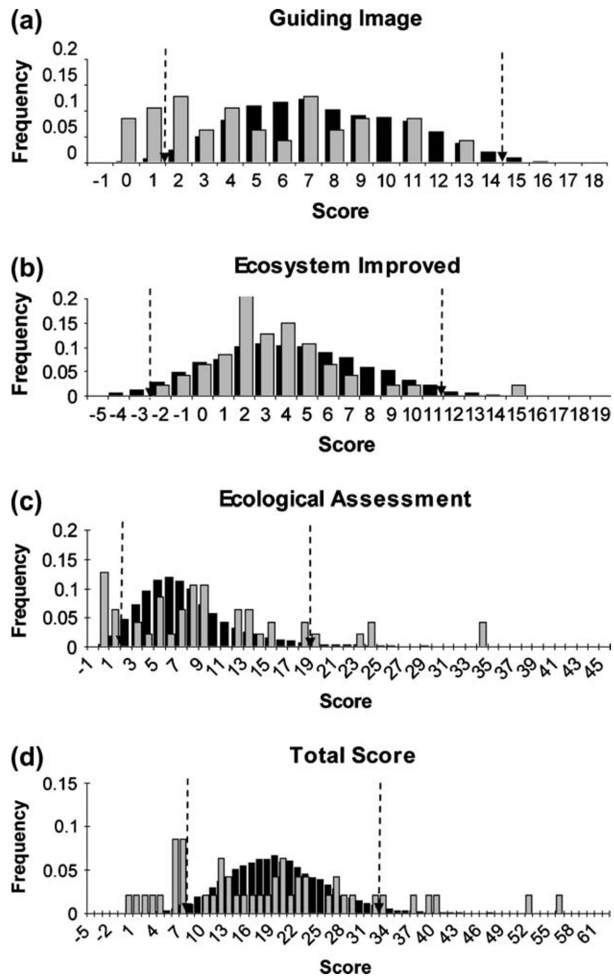

Figure 10. Results of numerical scoring assessment of survey categories: (a) Guiding Image, (b) Ecological Assessment, (c) Ecosystem Improved, and (d) Total Score. Data from phone surveys (in gray) are superimposed on the results of simulations (black). Ninety-five percent confidence intervals on the distribution are indicated by dashed vertical lines.

monitoring all of these projects, which would one choose to monitor? Complete monitoring designs, including the choice of monitored projects, are only determined by a fully articulated question. For those questions focused on specific projects, the choice of replicates is determined by the specific experimental design (Roni et al. 2002). In contrast, for questions focused at the watershed scale, all projects are monitored for implementation and this information is nested within a watershed scale, ecological assessment (Fausch et al. 2002). The intent of this article is not to advocate complex, detailed monitoring of every project. Rather it is hoped that local priorities revealed in interviewee responses may better inform future effectiveness monitoring designs at the appropriate scales.

Successful projects require an appreciation of the ecological complexities of diverse spatial and temporal scale process and linkages between large- and small-scale watershed processes (Lake 2001; Palmer et al. 2003; Wohl et al. 2005). As regional assessments continue to develop, the various components of project evaluation will similarly mature. In turn, this will help establish additional criteria for evaluating larger-scale coordination of projects. If monitoring is used to identify which types of actions work (or do not work) and under what circumstances planners can expect success in future restoration, then key elements of restoration design can be maintained through the prioritization, implementation, monitoring, and assessment phases of projects. Our survey demonstrates that in the PNW, there is still room for this maturation.

Conclusions

The interview process revealed the practitioners' view of stream restoration. Interviews detailed a wide diversity of definitions for monitoring, connectivity among projects, and project success. Given the lack of detailed and objective monitoring, it appears practitioners frequently rely on expert opinion and subjective impressions in assessing ecological impacts. Reasonable expectations for consistent effectiveness monitoring on the part of the project coordinators are limited by a lack of mandates and funding mechanisms that link monitoring to project execution. Establishing funding mechanisms to include long-term guarantees of financial support for extended effectiveness monitoring would be beneficial.

Monitoring is clearly a high priority for many reasons and should continue to be. Ideally, future inferences about project success would be advanced with a more consistent application of pre-treatment monitoring, measurements relevant to clearly articulated performance criteria, and meaningful coordination among projects and across large scales. When only 12% of respondents evaluate their success based on habitat or ecological measurements, critical opportunities to validate ecosystem response and maximize ecological benefits of regional funding are missed.

Implications for Practice

"Are there any lessons learned that you'd like to share with other practitioners?"

The following answers came directly from anonymous project respondents when asked what they have learned through their experience in restoration implementation:

- Project scale is significantly important! Do not limit yourself to the stream channel and focus as much or more on riparian area and floodplain connectivity.
- Do not take shortcuts—do it right or do not do it at all.
- It is great to use kids and schools to encourage community involvement.
- An enthusiastic landowner can be the key to success and failure. Although the implementation may be delayed, their support is crucial. In addition, the maintenance is taken care because of their willingness and support.
- Agencies and manuals can give us guidance, but we need personal points of view from farmers and ranchers. We need to maintain a personal relationship.
- Fencing needs a lot of maintenance. Landowners have the contract, but they may not necessarily perform the required maintenance work.
- If you do not have enough money to monitor the project, you do not have enough money to do the project.
- To eliminate deer/elk consumption of newly planted riparian areas, coplant spruce and cedar together in the same years, which allows them to establish above the browse line. After a couple of years, the spruce can be taken out once the cedar has been established.
- To improve leveled, flood-prone land containing invasive reed-canary grass, excavate (in dry season) mounds of dirt 2–3 ft tall. On the mounds, plant conifers. On the indentations, plant hard wood and shrubs tolerant to water. This will greatly increase survival rates and deter reed-canary grass, in addition to creating heterogeneous topographical habitat.

Acknowledgments

We would like to thank the interview respondents, without whom this project would not have been possible. We would also like to thank contributing editors and other NRRSS team members. The NRRSS working group is supported by the National Center for Ecological Analysis and Synthesis. The national effort received support from the National Science Foundation, David and Lucille Packard Foundation, Altria, U.S. Geological Survey, and the U.S. Environmental Protection Agency. The PNW regional team received support from the EPA, Bureau of Land Management, and Bureau of Reclamation, of which we would like to thank Michael Newsom. Additional support was provided by Congressional Award: Fund for the Improvement of Postsecondary Education (Award Number P116Z010107) Advanced Computing and Modeling Laboratory at the University of Idaho and the National Science Foundation Idaho EPSCoR Research Infrastructure Improvement Grant Grand Challenge Initiative, 2005–2008. The results, analysis, and conclusions do not represent the position or policy of the U.S. Department of Commerce, National Oceanic and Atmospheric Administration (NOAA), or NOAA Fisheries Service. More information on the NRRSS effort is available at www.restoringrivers.org.

LITERATURE CITED

Bash, J. S., and C. M. Ryan. 2002. Stream restoration and enhancement projects: is anyone monitoring? Environmental Management **29:**877–885.

Bernhardt, E. S., M. A. Palmer, J. D. Allen, G. Alexander, K. Barnas, S. Brooks, et al. 2005. Synthesizing U.S. river restoration efforts. Science **308:**636–637.

Bernhardt, E. S., E. B. Sudduth, M. A. Palmer, J. D. Allan, J. L. Meyer, G. Alexander, et al. 2007. Restoring rivers one reach at a time: results from a survey of U.S. river restoration practitioners. Restoration Ecology **15:**482–493.

Carpenter, S. R., T. M. Frost, D. Heisey, and T. K. Kratz. 1989. Randomized intervention analysis and the interpretation of whole-ecosystem experiments. Ecology **70:**1142–1152.

Downes, B. J., L. A. Barmuta, P. G. Fairweather, D. P. Faith, M. J. Keough, P. S. Lake, B. D. Mapstone, B. D. Quinn. 2002. Monitoring ecological impacts: concepts and practice in flowing waters. Cambridge University Press, New York.

Efron, B., and R. LaPage. 1992. Introduction to bootstrap. Pages 3–10 in R. LaPage and L. Billard, editors. Exploring the limits of the bootstrap. Wiley and Sons, New York.

Fausch, K. D., C. E. Torgersen, C. V. Baxter, and H. W. Li. 2002. Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes. Bioscience **52:**483–498.

Frissell, C. A., and S. C. Ralph. 1998. Stream and watershed restoration. Page 705 in R. J. Naiman and R. E. Bilby, editors. River ecology and management. Springer-Verlag, New York.

GAO (General Accounting Office). 2002. Columbia river basin salmon and steelhead: federal agencies' recovery responsibilities, expenditures, and actions. United States General Accounting Office Technical Report. GAO-02–612. United States General Accounting Officer, Washington, D.C.

Green, R. H. 1979. Sampling design and statistical methods for environmental biologists. New York.

Hall, P. 1992. Efficient bootstrap simulation. Pages 127–143 in R. LaPage and L. Billard, editors. Exploring the limits of the bootstrap. Wiley and Sons, New York, New York.

Katz, S., K. Barnas, V. R. Hicks, J. Cowen, and R. Jenkinson. 2007. Freshwater habitat restoration actions in the Pacific Northwest: a 10-year census. Restoration Ecology **15:**494–505.

Kondolf, G. M. 1998. Lessons learned from river restoration projects in California. Aquatic Conservation: Marine and Freshwater Ecosystem **8:**39–52.

Kondolf, G. M., and E. R. Micheli. 1995. Evaluating stream restoration projects. Environmental Management **19:**1–15.

Kondolf, G. M., M. W. Smeltzer, and S. Railsback. 2001. Design and performance of a channel reconstruction project in a coastal California gravel-bed stream. Environmental Management **28:**761–776.

Lake, P. S. 2001. On the maturing of restoration: linking ecological research and restoration. Ecological Management and Restoration **2:**110–115.

Lavendel, B. 2002. The business of ecological restoration. Ecological Restoration **20:**173–178.

MacDonald, L. H., A. W. Smart, and R. C. Wissmer. 1991. Monitoring guidelines to evaluate effects of forestry activities on streams in the Pacific Northwest and Alaska. U.S. Environmental Protection Agency. Seattle, Washington. EPA 910/9-91-001.

Malakoff, D. 2004. The river doctor. Science **305:**937–939.

NRC (National Research Council). 1992. Restoration of aquatic ecosystems: science, technology, and public policy. National Academy Press, Washington, D.C.

NRC (National Research Council). 1996. Upstream: salmon and society in the Pacific Northwest. National Academy Press, Washington, D.C.

Palmer, M. A., E. S. Bernhardt, J. D. Allen, G. Alexander, S. Brooks, S. Carr, et al. 2005. Standards for ecologically successful river restoration. Journal of Applied Ecology **42:**208–217.

Palmer, M. A., D. Hart, J. D. Allen, E. S. Bernhardt, and NRRSS Working Group. 2003. Bridging engineering, ecological, and geomorphic science to enhance riverine restoration: local and national efforts. Proceedings of a National Symposium on Urban and Rural Stream Protection and Restoration, EWRI World Water and Environmental Congress, Philadelphia, Pennsylvania, June 2003, American Society of Civil Engineers, Reston, Virginia.

Paulsen, C. M., and T. R. Fisher. 2005. Do habitat actions affect juvenile survival? An information-theoretic approach applied to endangered Snake River Chinook Salmon. Transaction of the American Fisheries Society **134:**68–85.

Reeves, G. H., F. H. Everest, and J. R. Sedell. 1991*a*. Responses of anadromous salmonids to habitat modifications: how do we measure them? Pages 62–67 in J. Colt and R. J. White, editors. Fisheries Bioengineering Symposium, American Fisheries Society, Symposium 10. Bethesda, Maryland.

Reeves, G. H., J. D. Hall, T. D. Roelofs, T. L. Hickman, and C. O. Baker. 1991*b*. Rehabilitating and modifying stream habitats. Pages 519–557 in W. R. Meehan, editor. Influences of forest and rangeland management on salmonid fishes and their habitats. Special Publications 19. American Fisheries Society, Bethesda, Maryland.

Roni, P., T. J. Beechie, R. E. Bilby, F. E. Leonetti, M. M. Pollock, and G. P. Pess. 2002. A review of stream restoration techniques and a hierarchical strategy for prioritizing restoration in Pacific Northwest watersheds. North American Journal of Fisheries Management **22:**1–20.

Rosenbaum, P. R. 2002. Observational Studies. 2nd edition. Springer-Verlag, New York.

Ruckelshaus, M. H., P. S. Levin, J. B. Johnson, and P. M. Kareiva. 2002. The Pacific salmon wars: what science brings to the challenge of recovering species. Annual Review Ecology and Systematics **33:**665–706.

Sutherland, W. J., A. S. Pullin, P. M. Dolman, and T. M. Knight. 2004. The need for evidence-based conservation. Trends in Ecology and Evolution **19:**305–308.

Thompson, W. L., G. C. White, and C. Gowan. 1998. Monitoring vertebrate populations. Academic Press, San Diego, California.

Underwood, A. J. 1994. Beyond BACI: sampling designs that might reliably detect environmental disturbances. Ecological Applications **4:**3–15.

Wohl, E., P. L. Angermeier, B. Bledsoe, G. M. Kondolf, L. MacDonnell, D. M. Merritt, M. A. Palmer, N. L. Poff, and D. Tarboton. 2005. River restoration. Water Resources Research 41, W10301, doi: 10.1029/2005WR003985.

Zar, J. H. 1984. Biostatistical Analysis. 2nd edition. Simon & Schuster Co, Englewood Cliffs, New Jersey.